Université de Nice Sophia-Antipolis
L2 MASS
TD de statistique

Année Universitaire 2014/2015
Statistique

Correction détaillée du QCM 1

1. Rappel des résultats du cours

Proposition 1 (Inégalité de Markov). *Soit X une v.a.* <u>*positive*</u>*, alors*

$$\forall t > 0, \quad \mathbb{P}\left(X \geq t\right) \leq \frac{\mathbb{E}\left[X\right]}{t}.$$

Proposition 2 (Inégalité de Bienaymé-Tchébitchef). *Soit X une v.a., alors*

$$\forall t > 0, \quad \mathbb{P}\left(\left|X - \mathbb{E}\left[X\right]\right| \geq t\right) \leq \frac{\mathbb{V}\mathrm{ar}\left[\mathrm{X}\right]}{t^2}.$$

Théorème 3 (Loi des grands nombres). *Soit $(X_i)_{i\geq 1}$ une suite de variables i.i.d. (indépendantes et de même loi) telle que $\mathbb{E}\left[X_1\right] < \infty$.*

Alors, on a $\bar{X}_n \xrightarrow[n\to\infty]{} m$ presque sûrement, où $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$.

Théorème 4 (Théorème Central Limite). *Soit $(X_i)_{i\geq 1}$ une suite de variables i.i.d. telle que $\mathbb{V}\mathrm{ar}\left[\mathrm{X}_1\right] < \infty$.*

Alors

$$\sqrt{n}\frac{\bar{X}_n - \mathbb{E}\left[X_1\right]}{\sqrt{\mathbb{V}\mathrm{ar}\left[\mathrm{X}_1\right]}} \xrightarrow[n\to\infty]{} \mathcal{N}(0,1) \quad \text{en loi .}$$

Théorème 5 (Lemme de Slutsky). *Soit $(Z_n)_{n\geq 1}$ une suite de v.a. convergeant en loi vers Z. Soit $(U_n)_{n\geq 1}$ une suite de v.a. qui converge presque sûrement vers 1.*

Alors, $\widetilde{Z}_n = U_n Z_n \xrightarrow[n\to\infty]{} Z$ en loi.

2. Correction

La correction est divisée en deux sections qui correspondent aux deux exercices. Chaque section est divisée en trois sous-sections correspondant aux trois question de chaque exercice. Les notations définies sur le sujet ne sont pas rappelées ici. Les résultats qui faisaient partis des bonnes réponses à cocher son soulignés.

Correction 1.

(1) La densité de X_1 est donnée par f, donc

$$\mathbb{E}\left[X_1\right] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^\theta x\theta^{-1} dx = \theta^{-1}\left[\frac{x^2}{2}\right]_0^\theta = \frac{\theta}{2}.$$

De même,

$$\mathbb{E}\left[X_1^2\right] = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^\theta x^2\theta^{-1} dx = \theta^{-1}\left[\frac{x^3}{3}\right]_0^\theta = \frac{\theta^2}{3}.$$

Ainsi, $\underline{\mathbb{E}\left[X_1\right] = \theta/2}$ et $\underline{\mathbb{V}\mathrm{ar}\left(\mathrm{X}_1\right) = \theta^2/3 - \theta^2/4 = \theta^2/12}$. Donc la question n'admettait qu'une seule bonne réponse, les autres étant fausses par contradiction.

Pour anticiper les questions suivantes, on peut dire que

$$\mathbb{E}\left[\bar{X}_n\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[X_i\right] = \mathbb{E}\left[X_1\right] = \frac{\theta}{2}$$

et

$$\mathbb{V}\mathrm{ar}\left(\bar{\mathrm{X}}_\mathrm{n}\right) = \mathbb{V}\mathrm{ar}\left(\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\mathrm{X}_\mathrm{i}\right) = \frac{1}{\mathrm{n}^2}\mathbb{V}\mathrm{ar}\left(\sum_{\mathrm{i}=1}^{\mathrm{n}}\mathrm{X}_\mathrm{i}\right) = \frac{1}{\mathrm{n}^2}\sum_{\mathrm{i}=1}^{\mathrm{n}}\mathbb{V}\mathrm{ar}\left(\mathrm{X}_\mathrm{i}\right) = \frac{1}{\mathrm{n}}\mathbb{V}\mathrm{ar}(\mathrm{X}_1) = \frac{\theta^2}{12\mathrm{n}},$$

où on utilise que la variance d'un somme de v.a. indépendantes est égale à la somme de ces v.a.

(2) Si on applique la Proposition 1 à X_1 qui est positive, cela donne :

$$\underline{\forall t > 0, \quad \mathbb{P}\left(X_1 \geq t\right) \leq \theta/(2t)},$$

car $\mathbb{E}\left[X_1\right] = \theta/2$.

Si on applique la Proposition 1 à $\bar{X}_n$ pour tout $n \geq 1$, cela donne :

$$\forall t > 0, \forall n \geq 1, \quad \mathbb{P}\left(\bar{X}_n \geq t\right) \leq \theta/(2t),$$

et non ce qui est proposé car $\mathbb{E}\left[\bar{X}_n\right] = \theta/2$.

Si on applique la Proposition 2 à $2\bar{X}_n$ pour tout $n \geq 1$, cela donne :

$$\underline{\forall a > 0, \forall n \geq 1, \quad \mathbb{P}\left(\left|2\bar{X}_n - \theta\right| \geq a\right) \leq \theta^2/(3na^2)},$$

car $\mathbb{V}\mathrm{ar}(2\bar{\mathrm{X}}_\mathrm{n}) = 4\mathbb{V}\mathrm{ar}(\bar{\mathrm{X}}_\mathrm{n}) = \theta^2/(3\mathrm{n})$. L'autre inégalité proposée est fausse. En effet, si on prend $n = 1$ et $a = \theta/2$, on aboutit à $\mathbb{P}\left(|2X_1 - \theta| \leq \theta/2\right) = 1/2 \leq 1/3$.

(3) Si on applique le Théorème 3 à la suite $(X_i)_{i\geq 1}$, cela donne :

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{} \mathbb{E}\left[X_1\right] = \frac{\theta}{2} \quad p.s.$$

En revanche, on ne peut rien en déduire sur la limite de $\lfloor\frac{4\bar{X}_n}{\theta}\rfloor$ parce que la fonction $x \mapsto \lfloor\frac{4x}{\theta}\rfloor$ n'est pas continue en $x = \theta/2$. Dans le cadre de l'exercice, l'énoncé est faux car $\forall n \geq 1$, $\mathbb{P}\left(\lfloor\frac{4\bar{X}_n}{\theta}\rfloor = 1\right) = 1/2$ et ne tend pas vers 0.

Si on applique le Théorème 3 à la suite $(X_i^2)_{i\geq 1}$, cela donne :

$$\underline{\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow[n\to\infty]{} \mathbb{E}\left[X_1^2\right] = \frac{\theta^2}{3} \quad p.s.}$$

Si on applique le Théorème 4 à la suite $(X_i)_{i\geq 1}$, cela donne :

$$Z_n = \sqrt{n}\frac{\bar{X}_n - \theta}{\theta/(\sqrt{12})} = \sqrt{3n}\frac{2\bar{X}_n - \theta}{\theta} \xrightarrow[n\to\infty]{} \mathcal{N}(0,1) \quad \text{en loi.}$$

En particulier, on peut en déduire que $\underline{\sqrt{3n}(2\bar{X}_n - \theta) \xrightarrow[n\to\infty]{} \mathcal{N}(0,\theta^2) \text{ en loi.}}$

Notons $U_n = \theta/2\bar{X}_n$ pour tout $n \geq 1$. La loi des grands nombres nous dit que $2\bar{X}_n/ \to \theta$ p.s. Remarquons que $\theta \neq 0$. Alors, par continuité de la fonction inverse en tout point différent de 0, on en déduit que $U_n \to 1$ p.s. Si on applique le Théorème 5 à la suite $(Z_n)_{n\geq 1}$ qui converge en loi vers $\mathcal{N}(0,1)$ et à la suite $(U_n)_{n\geq 1}$ qui converge p.s. vers 1, cela donne :

$$\widetilde{Z}_n = \sqrt{n}\frac{\bar{X}_n - \theta}{(2\bar{X}_n)/(\sqrt{12})} = \sqrt{3n}\left(1 - \frac{\theta}{\bar{X}_n}\right) \to \mathcal{N}(0,1) \quad \text{en loi.}$$

En particulier, par définition de la convergence en loi, on peut en déduire que

$$\underline{\forall a < b, \quad \lim_{n\to\infty} \mathbb{P}\left(a \le \sqrt{3n}\left(1 - \frac{\theta}{\bar{X}_n}\right) \le b\right) = \mathbb{P}\left(a \le \mathcal{N}(0,1) \le b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx.}$$

Correction 2.

(1) La loi de X_1 est donnée par $\mathbb{P}(X_1 = -1) = \mathbb{P}(X_1 = 0) = 1/4$ et $\mathbb{P}(X_1 = 1) = 1/2$, donc

$$\mathbb{E}[X_1] = (-1).(1/4) + (0).(1/4) + (1).(1/2) = 1/4.$$

De même,

$$\mathbb{E}\left[X_1^2\right] = (-1)^2.(1/4) + (0)^2.(1/4) + (1)^2.(1/2) = 3/4.$$

Ainsi, $\underline{\mathbb{E}[X_1] = 1/4}$ et $\underline{\mathbb{V}\mathrm{ar}(\mathrm{X}_1) = 3/4 - 1/16 = 11/16}$. De plus, on peut dire que

$$\underline{\mathbb{E}\left[\bar{X}_n\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] = \mathbb{E}[X_1] = 1/4}$$

et

$$\mathbb{V}\mathrm{ar}\left(\bar{\mathrm{X}}_\mathrm{n}\right) = \mathbb{V}\mathrm{ar}\left(\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{X}_\mathrm{i}\right) = \frac{1}{\mathrm{n}^2}\mathbb{V}\mathrm{ar}\left(\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{X}_\mathrm{i}\right) = \frac{1}{\mathrm{n}^2}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathbb{V}\mathrm{ar}(\mathrm{X}_\mathrm{i}) = \frac{1}{\mathrm{n}}\mathbb{V}\mathrm{ar}(\mathrm{X}_1) = \frac{11}{16\mathrm{n}},$$

où on utilise que la variance d'un somme de v.a. indépendantes est égale à la somme de ces v.a.

(2) On ne peut pas appliquer la Proposition 1 à X_1 car ce n'est pas une variable positive ($\mathbb{P}(X_1 = -1) = 1/4 \neq 0$). De plus, l'inégalité proposée est fausse comme on peut le voir en prenant $t = 1$ qui donne $\mathbb{P}(X_1 \ge 1) = 1/2 \le 1/4$. En revanche, on peut appliquer la Proposition 1 à $\bar{X}_n + 1$, cela donne :

$$\underline{\forall t > 0, \quad \mathbb{P}\left(\bar{X}_n + 1 \ge t\right) \le 5/(4t)},$$

car $\mathbb{E}\left[\bar{X}_n + 1\right] = 5/4$.

Si on applique la Proposition 2 à $4\bar{X}_n$ pour tout $n \ge 1$, cela donne :

$$\underline{\forall a > 0, \forall n \ge 1, \quad \mathbb{P}\left(\left|4\bar{X}_n - 1\right| \ge a\right) \le 11/(na^2)},$$

car $\mathbb{V}\mathrm{ar}(4\bar{\mathrm{X}}_\mathrm{n}) = 16\mathbb{V}\mathrm{ar}(\bar{\mathrm{X}}_\mathrm{n}) = 11$. L'autre inégalité proposée est fausse. En effet, si on prend $n = 1$ et $a = 3$, on aboutit à $\mathbb{P}(|4X_1 - 1| \le 3) = 3/4 \le 11/36$.

(3) Si on applique le Théorème 3 à la suite $(X_i)_{i\ge1}$, cela donne :

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \xrightarrow[n\to\infty]{} \mathbb{E}[X_1] = \frac{1}{4} \quad p.s.$$

La fonction $x \mapsto x^2$ est continue en $x = 1/4$, donc on en déduit que $\underline{(\bar{X}_n)^2 \to 1/16}$ p.s.

Si on applique le Théorème 3 à la suite $(X_i^2)_{i\ge1}$, cela donne :

$$\frac{1}{n}\sum_{i=1}^n X_i^2 \xrightarrow[n\to\infty]{} \mathbb{E}\left[X_1^2\right] = 3/4 \neq 1/2 \quad p.s.$$

Si on applique le Théorème 4 à la suite $(X_i)_{i\ge1}$, cela donne :

$$Z_n = \sqrt{n}\frac{\bar{X}_n - (1/4)}{\sqrt{11/16}} = \sqrt{n}\frac{\bar{X}_n - (1/4)}{\sqrt{11}/4} \xrightarrow[n\to\infty]{} \mathcal{N}(0,1) \quad \text{en loi.}$$

En particulier, on peut en déduire que $\sqrt{n}(4\bar{X}_n - 1) \xrightarrow[n\to\infty]{} \mathcal{N}(0, 11)$ en loi. Pour finir, par définition de la convergence en loi, on peut en déduire que

$$\forall a < b, \quad \lim_{n\to\infty} \mathbb{P}\left(a \leq \sqrt{n}\left(4\bar{X}_n - 1\right) \leq b\right) = \mathbb{P}\left(a \leq \mathcal{N}(0, 11) \leq b\right) = \int_a^b \frac{1}{\sqrt{22\pi}} e^{-\frac{x^2}{22}}\, dx.$$